RELATIONSHIP BETWEEN COST OVERRUNS AND COMPLEXITY IN ENGINEERING PROJECTS: A MIXED APPROACH

BOHÓRQUEZ-CASTELLANOS, Jherson (1); MEJÍA, Guillermo (2)

(1) Universidad Industrial de Santander, (57) 3145141782, e-mail: jhersonbohorquez@gmail.com
(2) Universidad Industrial de Santander, e-mail: gmejia@uis.edu.co

ABSTRACT

Currently, cost overrun is a worldwide phenomenon in engineering projects. Several studies have quantified cost overruns in projects, while others have determined the leading factors of cost overruns. Among these factors, researchers and practitioners have taken an interest in construction complexity. Some authors suggest further analysis of the relationship between project complexity and project success because construction complexity influences decision-making and produces adverse effects on project cost performance. This study aimed at both an alternative mixed analysis of project complexity and cluster analysis of cost performance, based on project data of published studies. This study proposed three levels of complexity, validated by a clustering silhouette coefficient (> 0.7), which indicated the extent of cohesion among projects. Projects related to urban infrastructure development indicated low complexity; projects related to transport infrastructure projects, medium complexity; and projects related to special construction, involving high uncertainty, indicated high complexity. The computed pooled mean of cost overruns suggested that projects with high complexity are prone to have higher cost overruns (25.6%) on average; low complexity projects showed a cost overrun of 8.0% on average, and medium complexity projects showed a cost overrun of 9.0% on average. This evidence leads practitioners to be aware of the complexity and its relationship with cost overruns.

Keywords: Cost overruns, projects complexity, cluster analysis, confidence intervals

1 INTRODUCTION

Currently, cost overrun is a worldwide phenomenon in engineering projects (LOVE et al., 2017) that has been studied during the last decades. Several studies have quantified cost overruns in construction and engineering projects (AL-HAZIM; SALEM; AHMAD, 2017; SHRESTHA; BURNS; SHIELDS, 2013; ODECK, 2004; FLYVBJERG; SKAMRIS; BUHL, 2003). Most of them were based on historical data, where the information quality is crucial to assess the cost performance of executed projects. The information of these cost overruns and their final log depends on: first, the data reliability of the project’s stakeholder (i.e., owners, contractors, designers, consultants); and secondly, the formalization level of the lessons learned that have been established and undertook in each project.
On the other hand, other studies have determined the leading factors of cost overruns (MEMON; RAHMAN, 2013; PARK; PAPADOPOULOU, 2012; MANSFIELD; UGWU; DORAN, 1994). Among these factors, researchers and practitioners have taken an interest in construction complexity (NGUYEN et al., 2019). Some authors advocate analyzing construction complexity associated with the management process and not with the physical characteristics of the projects, which implies mostly a qualitative analysis (KIAN MANESH RAD; SUN; BOSCHÉ, 2017). Other authors suggest that construction complexity should not be interpreted as risk management, though it is often associated with uncertainty levels of the project (DAO et al., 2016). Hence, a definition of project complexity must include several viewpoints, from multiple disciplines involved in the engineering projects. According to Dao’s research, “Project complexity is the degree of differentiation of project elements, interrelatedness between project elements, and consequential impact on project decisions.”

Project complexity influences decision-making and produces adverse effects on project cost performance. Therefore, some authors suggest further analysis of the relationship between project complexity and project success (LUO et al., 2017) through integrating qualitative and quantitative research methods (DAMAYANTI; HARTONO; WIJAYA, 2019). Thus, this study aimed at both an alternative mixed analysis of the complexity of engineering projects and cluster analysis of cost performance, based on project data of published studies.

2 LITERATURE REVIEW

Recently studies have suggested that project complexity is affecting project success (NGUYEN et al., 2019; DAMAYANTI; HARTONO; WIJAYA, 2019; MIRZA; EHSAN, 2017; DE CARVALHO; PATAH; DE SOUZA BIDO, 2015). These authors are investigating and associating complexity as a factor (NGUYEN et al., 2019; MAJEROWICZ; SHINN, 2016; KARDES et al., 2013) that influences the cost and schedule performance. Nonetheless, most of these studies show no consensus on project complexity definition (MIRZA; EHSAN, 2017), which affects the practice of researching this field.

Some researchers have used qualitative methods to understand and explain complexity by disaggregating it in either components, elements, or classifications, based on theoretical approaches (KIAN MANESH RAD; SUN; BOSCHÉ, 2017). In fact, in megaprojects, Damayanti, Hartono and Wijaya (2019) classified project complexity as structural, emergent, and social-politic complexity. Kian Manesh Rad, Sun and Bosché (2017) classified complexity as descriptive complexity (i.e., related to the intrinsic property of the project) and perceived complexity. Kardes et al. (2013) classified complexity in terms of technical and social complexity. These different classification viewpoints represent a challenge in studying project complexity.

Therefore, for an initial understanding of construction complexity, the relationship between complexity and project performance must be established. Although this relationship has been widely perceived through qualitative analyses, a quantitative analysis must be addressed. In terms of using quantitative data, Nguyen et al. (2019) found a significant correlation between project complexity and schedule growth, but not with cost growth;
some limitations expressed by these authors are referred to the sample size used (i.e.,
empirical data from 79 transportation projects) and the individual perceptions influencing
the data used to compute cost growth.

According to trends and challenges in project complexity research, this study performed
both a quantitative analysis based on the cost overruns aggregation for three groups of
complexity, and a cluster analysis that qualitatively evaluated these groups, which allowed
the evaluation between groups through confidence intervals for the average cost overrun of
each group.

3 METHODOLOGY
This research used a mixed approach to achieve this study’s aim. A systematic review of
the literature collected cost overrun data and analyzed the complexity of the construction
projects by clustering such overrun data. The systematic reviews aid researchers in
discovering patterns and trends from primary research publications. Evidence found in
primary studies is assessed, summarized, and interpreted under methodological procedures
(BORREGO; FOSTER; FROYD, 2014; KHAN et al., 2003).

3.1 Data collection of projects cost overruns
The systematic review of this study used searching criteria and Boolean equations to finally
gather 22 primary articles and cost data of 2,598 construction projects (see Table 1). The
searching process used keywords such as “cost overrun”, “cost underrun”, “cost deviation”,
“cost growth”, “construction projects”, and “engineering projects”, in the following
databases: ASCE, EBSCO, Emerald, Science Direct, Springer, Taylor & Francis, and Web
of Science. This searching was delimited to publication years between 1985 and 2018. The
primary evidence included statistics reported in each study, such as sample size, the mean
and standard deviation of projects’ cost performance.

The assessment of the primary studies began with the classification of cost overruns by
geographical region of the project. Each region was summarized by basic statistics of its
projects (i.e., number of projects, the mean and standard deviation of cost overruns). For
this research and under the assumption of a well-defined project’s scope, those projects
with cost overruns higher than 100% were excluded. Projects well planned with a clear
scope definition should not exceed established tolerance levels. According to AACE’s
recommended practices, for an estimate of level 5 (i.e., concept screening of project), cost
estimates deviations at least of 100% are expected for those projects with a level of scope
definition between 0% and 2% (HAMILTON, 2004).

Table 1 – Articles analyzed

<table>
<thead>
<tr>
<th>Article</th>
<th>Projects analyzed</th>
<th>Projects Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>(KOCH, 2012)</td>
<td>10</td>
<td>Europe</td>
</tr>
<tr>
<td>(LANGFORD et al., 2003)</td>
<td>11</td>
<td>Europe</td>
</tr>
<tr>
<td>(LIU; NAPIER, 2010)</td>
<td>30</td>
<td>Oceania</td>
</tr>
<tr>
<td>(LORENTZEN; OGLEND; OSMUNDSEN, 2017)</td>
<td>79</td>
<td>Europe</td>
</tr>
<tr>
<td>Source/Year</td>
<td>Country</td>
<td>Projects Count</td>
</tr>
<tr>
<td>------------------------------------------------</td>
<td>----------------------------------</td>
<td>----------------</td>
</tr>
<tr>
<td>SOVACOOL; GILBERT; NUGENT, 2014</td>
<td>North America (91), Latin America (17), Africa (15), Asia (76), Europe (129), Oceania (12)</td>
<td></td>
</tr>
<tr>
<td>KALIBA; MUYA; MUMBA, 2009</td>
<td>Africa</td>
<td>7</td>
</tr>
<tr>
<td>SHEHU et al., 2014</td>
<td>Asia</td>
<td>139</td>
</tr>
<tr>
<td>FORCADA et al., 2014</td>
<td>Europe</td>
<td>8</td>
</tr>
<tr>
<td>SHRESTHA; BURNS; SHIELDS, 2013</td>
<td>North America</td>
<td>363</td>
</tr>
<tr>
<td>LOVE et al., 2010</td>
<td>Oceania</td>
<td>115</td>
</tr>
<tr>
<td>NASSAR; NASSAR; HEGAB, 2005</td>
<td>North America</td>
<td>219</td>
</tr>
<tr>
<td>CHEN et al., 2015</td>
<td>North America</td>
<td>101</td>
</tr>
<tr>
<td>RWAKAREHE; MFINANGA, 2013</td>
<td>Africa</td>
<td>7</td>
</tr>
<tr>
<td>AL-HAZIM; SALEM; AHMAD, 2017</td>
<td>Asia</td>
<td>9</td>
</tr>
<tr>
<td>RAMANATHAN; SAMBU POTTY; BIN IDRUS, 2011</td>
<td>Asia</td>
<td>7</td>
</tr>
<tr>
<td>AWOJOBI; JENKINS, 2016</td>
<td>Africa (13), Latin America (15), Asia (22), Europe (5), Oceania (3)</td>
<td></td>
</tr>
<tr>
<td>VALLEJO-BORDA et al., 2015</td>
<td>Latin America</td>
<td>109</td>
</tr>
<tr>
<td>MAKOVŠEK; TOMINC; LOGOŽAR, 2012</td>
<td>Europe</td>
<td>33</td>
</tr>
<tr>
<td>CANTARELLI et al., 2012</td>
<td>Europe</td>
<td>78</td>
</tr>
<tr>
<td>ODECK, 2004</td>
<td>Europe</td>
<td>620</td>
</tr>
<tr>
<td>FLYVBJERG; SKAMRIS; BUHL, 2003</td>
<td>Europe (181), North America (61)</td>
<td></td>
</tr>
<tr>
<td>LOVE et al., 2017</td>
<td>Oceania</td>
<td>13</td>
</tr>
</tbody>
</table>

Source: Authors

### 3.2 Relation of project complexity with cost overruns

Based on two basic statistics (i.e., mean and standard deviation -SD), the projects of the sample were qualitatively classified into three groups (see Image 1). This classification allowed the differentiation of the construction complexity level. The grouping criteria were: projects with shared characteristics, project type, and order of magnitude of cost overruns. Once the groups were established, a cluster analysis was done to see the reliability and consistency of data in the arranged groups. This cluster analysis validated the initial proposal of project groups.

Cluster analysis was performed in “Orange”, an open-source machine learning and data visualization, developed by the University of Ljubljana. The study used a hierarchical clustering controlled by the extent of consistency within each cluster. Then, the data were visualized using the silhouette plot (see Image 3). Hierarchical clustering was performed under the assumption of the presence of the three complexity groups shown in Image 1. Thus, the silhouette coefficient of the grouped data was computed to indicate the clustering quality of data: “A silhouette value of more than 0.7 is considered as a strong cluster value” (ROUSSEEUW, 1987; ZHAO et al., 2018).
Finally, an aggregated analysis of cost overruns by complexity allowed the inference into the population. With confidence intervals at 95%, the study evaluated the cost overrun differences across the three complexity groups.

4 RESULTS

4.1 Cost overruns data

The 2,598 projects drawn from the 22 articles were categorized into six continental regions. Table 2 shows aggregated data by region according to the data gathered in each article. From an overall context, 2,598 engineering projects showed a pooled average mean of 11.5% of cost overrun, with a pooled standard deviation of 25.4%.

<table>
<thead>
<tr>
<th>Region</th>
<th>N Adjusted</th>
<th>Mean Pooled</th>
<th>SD Pooled</th>
</tr>
</thead>
<tbody>
<tr>
<td>Africa</td>
<td>42</td>
<td>23.9%</td>
<td>28.5%</td>
</tr>
<tr>
<td>Asia</td>
<td>253</td>
<td>8.2%</td>
<td>23.5%</td>
</tr>
<tr>
<td>Europe</td>
<td>1,154</td>
<td>14.3%</td>
<td>30.7%</td>
</tr>
<tr>
<td>Latin America</td>
<td>141</td>
<td>17.9%</td>
<td>29.5%</td>
</tr>
<tr>
<td>North America</td>
<td>835</td>
<td>6.7%</td>
<td>17.6%</td>
</tr>
<tr>
<td>Oceania</td>
<td>173</td>
<td>11.4%</td>
<td>14.1%</td>
</tr>
<tr>
<td>Overall</td>
<td>2,598</td>
<td>11.5%</td>
<td>25.4%</td>
</tr>
</tbody>
</table>

Source: Authors

4.2 Qualitative relationship between cost overruns and project complexity

Image 1 below shows the three complexity groups proposed from the collected data. The order of the groups indicates the level of complexity associated. On the left side, EPT3-group represents the projects of less complexity: utilities, facilities, industrial, and energy transmission projects. In the middle, EPT1-group represents the medium complexity projects: transport infrastructure projects, mainly roads and highways. Finally, on the right side, EPT2-group represents the projects with high complexity: bridges, tunnels, subways, rails, and energy generation projects. This complexity analyzed herein is associated mainly with the construction process of the project and its management process.

4.3 Quantitative relationship between cost overruns and project complexity

A hierarchical clustering approach associated the different classes of engineering projects, considering their mean and standard deviation of cost overrun (see Image 2). This association showed clustering patterns within the intrinsic behavior of data. The silhouette coefficient computed for this study was = 0.7, which indicates a strong cluster value, validating the consistency of complexity groups proposed (see Image 3).
Image 1 – Aggregation of project types by complexity

Engineering Projects Type (EPT)

- EPT3
  - Solar facility
  - Thermal plant
  - Transmission
  - Wind farm
  - Water infrastructure
  - Water main constructions
  - Water treatment plants
  - Water/Wastewater
  - Flood control
  - Industrial waste plants
  - Oil and Gas
  - Pumping stations
  - Sewer treatment plants
  - Subdivision development
  - Utilities

- EPT1
  - Highways
  - Infrastructure
  - Roads
  - Asphalt paving
  - Motorway construction
  - Runways
  - Transit & Transport infrastructure

- EPT2
  - Bridge
  - Fixed links (Bridges & Tunnels)
  - Rail
  - Nuclear reactor
  - Tunnel and subways
  - Wharves
  - Wind Power Plants (Offshore)
  - Airport
  - Elevated highways
  - Marinas
  - Reservoirs and dams
  - Sea walls

Source: Authors

Image 2 – Hierarchical clustering by project complexity

Source: Authors in Orange3
Based on an inferential analysis, this research found that the expected cost overruns in engineering projects will range between 10.5% and 12.4% (see Table 3). The cost overrun of the EPT2 group was 25.57%, i.e., 3.19 times greater than that of the EPT3 group, and 2.85 times greater than that of the EPT1 group. Although the sample size has a proportional influence on the width of the confidence intervals, the confidence intervals herein computed might suggest that projects on EPT2 have higher cost overruns than the other groups (see Image 4).

Table 3 – Aggregation of cost overruns by the complexity

<table>
<thead>
<tr>
<th>Project Type</th>
<th>N Adjusted</th>
<th>Mean Pooled</th>
<th>SD Pooled</th>
<th>CI 95% UL</th>
<th>CI 95% LL</th>
</tr>
</thead>
<tbody>
<tr>
<td>EPT1</td>
<td>1,684</td>
<td>8.97%</td>
<td>24.01%</td>
<td>10.1%</td>
<td>7.8%</td>
</tr>
<tr>
<td>EPT2</td>
<td>419</td>
<td>25.57%</td>
<td>33.28%</td>
<td>28.8%</td>
<td>22.4%</td>
</tr>
<tr>
<td>EPT3</td>
<td>495</td>
<td>8.00%</td>
<td>23.79%</td>
<td>10.1%</td>
<td>5.9%</td>
</tr>
<tr>
<td>Overall</td>
<td>2,598</td>
<td>11.46%</td>
<td>25.69%</td>
<td>12.4%</td>
<td>10.5%</td>
</tr>
</tbody>
</table>

Source: Authors
5 DISCUSSION AND CONCLUSIONS

This study adopted a mixed approach to understand complexity in engineering projects. According to this approach, the hierarchical clustering of the mean and standard deviation of cost overrun of aggregated projects, drawn from different and worldwide studies, indicated that EPT2 is the group of high complexity.

While the EPT3 group, related urban infrastructure development projects, indicated low complexity, the EPT1 group, related transport infrastructure projects, indicated medium complexity, and the EPT2 group, related projects involving high uncertainty, indicated high complexity. This classification agrees with the classification proposed by Santana (1990), in which the EPT2 group should be assigned to complex and singular project category (i.e., nuclear power stations, airports, tunnels, and dams).

One of the main findings of this research was the mixed approach used to identify, cluster, and validate the complexity level of engineering projects of reported basic statistics data. The cluster analysis identified and validated patterns of cost overruns among engineering projects. High consistency within proposed complexity groups of engineering projects, with a silhouette average value greater than 0.7 showed a correlated cost performance.

Engineering projects are prone to have (on average) greater cost overruns if their level of complexity increases. The aggregation of cost overruns for each group suggested that those engineering projects, belonging to the EPT2 group, are affected by higher cost overruns in comparison to the other groups. Further quantitative analysis with other statistical tests should show us if the complexity levels affect the cost performance of projects.
Based on this evidence, project managers must be aware of cost overruns and its relationship to complexity. Therefore, more efforts to plan and control construction processes must be undertaken when the level of complexity of projects is high. An effective decision-making process could mitigate the impact of cost overruns on project success if the construction managers know the level of complexity of their projects.

REFERENCES


Elsevier Ltd.


ACKNOWLEDGMENTS

The author Jherson Bohórquez thanks the Department of Santander in Colombia for the scholarship granted through Call 771 of the Administrative Department of Science, Technology, and Innovation (Colciencias).